



1 / 34

SEQUENCE LISTING

<110> NATIONAL RESEARCH COUNCIL OF CANADA

SU, Zhengding

NI, Feng

<120> Novel Fusion Proteins for the Production
of Recombinant Peptides

<130> 2139-22US

<150> US 60/402,075

<151> 2002-08-09

<160> 48

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 25

<212> PRT

<213> Staphylococcus nuclease

<400> 1

Ala Thr Ser Thr Lys Lys Leu His Lys Glu Pro Ala Thr Leu Ile Lys

1 5 10 15
Ala Ile Asp Gly Asp Thr Val Lys Leu
 20 25

<210> 2

<211> 5

<212> PRT

<213> Staphylococcus nuclease

<400> 2

Tyr Lys Gly Gln Pro
1 5

<210> 3

<211> 29

<212> PRT

<213> Staphylococcus nuclease

<400> 3

Leu Leu Leu Val Asp Thr Pro Glu Thr Lys His Pro Lys Lys Gly Val
1 5 10 15
Glu Lys Tyr Gly Pro Glu Ala Ser Ala Phe Thr Lys Lys
 20 25

<210> 4

<211> 32

<212> PRT

<213> Staphylococcus nuclease

<400> 4

Val Glu Asn Ala Lys Lys Ile Glu Val Glu Phe Asp Lys Gly Gln Arg

1 5 10 15

Thr Asp Lys Tyr Gly Arg Gly Leu Ala Tyr Ile Tyr Ala Asp Gly Lys

20 25 30

<210> 5

<211> 5

<212> PRT

<213> Staphylococcus nuclease

<400> 5

Val Asn Glu Ala Leu

1 5

<210> 6

<211> 6

<212> PRT

<213> Staphylococcus nuclease

<400> 6

Phe Asn Pro Arg Gly Ser

1 5

<210> 7

<211> 14

<212> PRT

<213> Staphylococcus nuclease

<400> 7

Val Arg Gln Gly Leu Ala Lys Val Ala Tyr Val Tyr Lys Pro

1 5 10

<210> 8

<211> 32

<212> PRT

<213> Staphylococcus nuclease

<400> 8

Asn Asn Thr His Glu Gln Leu Leu Arg Lys Ser Glu Ala Gln Ala Lys

1 5 10 15

Lys Glu Lys Leu Asn Ile Trp Ser Glu Asp Asn Ala Asp Ser Gly Gln

20 25 30

<210> 9

<211> 4

<212> PRT

<213> Artificial Sequence

<220>

<223> Peptidic Cleavage Site found on the fusion carrier protein derived from
Staphylococcus nuclease

<400> 9

Phe Asn Pro Arg

1

<210> 10

<211> 4

<212> PRT

<213> Artificial Sequence

<220>

<223> Peptidic Cleavage Site found on the fusion carrier protein derived from
Staphylococcus nuclease

<400> 10

Leu Val Pro Arg

1

<210> 11

<400> 11

000

<210> 12

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> Peptidic Cleavage Site found on the fusion carrier protein derived from *Staphylococcus nuclease*

<400> 12

Asp Asp Asp Asp Lys

1 5

<210> 13

<211> 351

<212> DNA

<213> Artificial Sequence

<220>

<223> Fusion Carrier Protein named SFC120

<220>

<221> CDS

<222> (1)...(351)

<400> 13

gca act tca act aaa aaa tta cat aaa gaa cct gcg act tta att aaa 48

Ala Thr Ser Thr Lys Lys Leu His Lys Glu Pro Ala Thr Leu Ile Lys

1 5 10 15

gcg att gat ggt gat acg gtt aaa tta atg tac aaa ggt caa cca atg 96

Ala Ile Asp Gly Asp Thr Val Lys Leu Met Tyr Lys Gly Gln Pro Met

20

25

30

aca ttc aga cta tta ttg gtt gat aca cct gaa aca aag cat cct aaa 144

Thr Phe Arg Leu Leu Leu Val Asp Thr Pro Glu Thr Lys His Pro Lys

35

40

45

aaa ggt gta gag aaa tat ggt cct gaa gca agt gca ttt acg aaa aaa 192

Lys Gly Val Glu Lys Tyr Gly Pro Glu Ala Ser Ala Phe Thr Lys Lys

50

55

60

atg gta gaa aat gca aag aaa att gaa gtc gag ttt gac aaa ggt caa 240

Met Val Glu Asn Ala Lys Lys Ile Glu Val Glu Phe Asp Lys Gly Gln

65

70

75

80

aga act gat aaa tat gga cgt ggc tta gcg tat att tat gct gat gga 288

Arg Thr Asp Lys Tyr Gly Arg Gly Leu Ala Tyr Ile Tyr Ala Asp Gly

85

90

95

aaa atg gta aac gaa gct tta gtt cgt caa ggc ttg gct aaa gtt gct 336

Lys Met Val Asn Glu Ala Leu Val Arg Gln Gly Leu Ala Lys Val Ala

100

105

110

tat gtt tac aaa cct 351

Tyr Val Tyr Lys Pro

115

<210> 14

<211> 117

<212> PRT

<213> Artificial Sequence

<220>

<223> Fusion Carrier Protein named SFC120

<400> 14

Ala Thr Ser Thr Lys Lys Leu His Lys Glu Pro Ala Thr Leu Ile Lys

1 5 10 15

Ala Ile Asp Gly Asp Thr Val Lys Leu Met Tyr Lys Gly Gln Pro Met

20 25 30

Thr Phe Arg Leu Leu Leu Val Asp Thr Pro Glu Thr Lys His Pro Lys

35 40 45

Lys Gly Val Glu Lys Tyr Gly Pro Glu Ala Ser Ala Phe Thr Lys Lys

50 55 60

Met Val Glu Asn Ala Lys Lys Ile Glu Val Glu Phe Asp Lys Gly Gln

65 70 75 80

Arg Thr Asp Lys Tyr Gly Arg Gly Leu Ala Tyr Ile Tyr Ala Asp Gly

85 90 95

Lys Met Val Asn Glu Ala Leu Val Arg Gln Gly Leu Ala Lys Val Ala

100 105 110

Tyr Val Tyr Lys Pro

115

•

3

a

q

a

11

cat cct aaa aaa ggt gta gag aaa tat ggt cct gaa gca agt gca ttt 192

His Pro Lys Lys Gly Val Glu Lys Tyr Gly Pro Glu Ala Ser Ala Phe

50

55

60

acg aaa aaa atg gta gaa aat gca aag aaa att gaa gtc gag ttt gac 240

Thr Lys Lys Met Val Glu Asn Ala Lys Lys Ile Glu Val Glu Phe Asp

65

70

75

80

aaa ggt caa aga act gat aaa tat gga cgt ggc tta gcg tat att tat 288

Lys Gly Gln Arg Thr Asp Lys Tyr Gly Arg Gly Leu Ala Tyr Ile Tyr

85

90

95

gct gat gga aaa atg gta aac gaa gct tta gtt cgt caa ggc ttg gct 336

Ala Asp Gly Lys Met Val Asn Glu Ala Leu Val Arg Gln Gly Leu Ala

100

105

110

aaa gtt gct tat gtt tac aaa cct gaa ttc atg 369

Lys Val Ala Tyr Val Tyr Lys Pro Glu Phe Met

115

120

<210> 16

<211> 123

<212> PRT

<213> Artificial Sequence

<220>

9

2

<223> bacterial expression vector pHSN-Met65Leu that encodes a fusion protein, containing a sequence for a gene encoding the N-terminal nucleotide binding domain of *staph. nuclease*, designated SFC120

<220>

<221> CDS

<222> (1)...(393)

<400> 17

aca gcc atg gga cac cat cac cat cac cat ggc gca act tca act aaa 48

Thr Ala Met Gly His His His His His His Gly Ala Thr Ser Thr Lys

1

5

10

15

aaa tta cat aaa gaa cct gcg act tta att aaa gcg att gat ggt gat 96

Lys Leu His Lys Glu Pro Ala Thr Leu Ile Lys Ala Ile Asp Gly Asp

20

25

30

acg gtt aaa tta atg tac aaa ggt caa cca atg aca ttc aga cta tta 144

Thr Val Lys Leu Met Tyr Lys Gly Gln Pro Met Thr Phe Arg Leu Leu

35

40

45

ttg gtt gat aca cct gaa aca aag cat cct aaa aaa ggt gta gag aaa 192

Leu Val Asp Thr Pro Glu Thr Lys His Pro Lys Lys Gly Val Glu Lys

50

55

60

tat ggt cct gaa gca agt gca ttt acg aaa aaa ttg gta gaa aat gca 240

Tyr Gly Pro Glu Ala Ser Ala Phe Thr Lys Lys Leu Val Glu Asn Ala

65

70

75

80

aag aaa att gaa gtc gag ttt gac aaa ggt caa aga act gat aaa tat 288

Lys Lys Ile Glu Val Glu Phe Asp Lys Gly Gln Arg Thr Asp Lys Tyr

85

90

95

gga cgt ggc tta gcg tat att tat gct gat gga aaa atg gta aac gaa 336

Gly Arg Gly Leu Ala Tyr Ile Tyr Ala Asp Gly Lys Met Val Asn Glu

100

105

110

gct tta gtt cgt caa ggc ttg gct aaa gtt gct tat gtt tac aaa cct 384

Ala Leu Val Arg Gln Gly Leu Ala Lys Val Ala Tyr Val Tyr Lys Pro

115

120

125

gaa ttc atg

393

Glu Phe Met

130

<210> 18

<211> 131

<212> PRT

<213> Artificial Sequence

<220>

<223> fusion protein encoded by bacterial expression vector pHSN-Met65Leu, containing a sequence for a gene encoding the N-terminal nucleotide binding domain of *staph.* nuclease, designated SFC120

<400> 18

Thr Ala Met Gly His His His His His His Gly Ala Thr Ser Thr Lys

1 5 10 15

Lys Leu His Lys Glu Pro Ala Thr Leu Ile Lys Ala Ile Asp Gly Asp

20 25 30

Thr Val Lys Leu Met Tyr Lys Gly Gln Pro Met Thr Phe Arg Leu Leu

35 40 45

Leu Val Asp Thr Pro Glu Thr Lys His Pro Lys Lys Gly Val Glu Lys

50 55 60

Tyr Gly Pro Glu Ala Ser Ala Phe Thr Lys Lys Leu Val Glu Asn Ala

65 70 75 80

Lys Lys Ile Glu Val Glu Phe Asp Lys Gly Gln Arg Thr Asp Lys Tyr

85 90 95

Gly Arg Gly Leu Ala Tyr Ile Tyr Ala Asp Gly Lys Met Val Asn Glu

100 105 110

Ala Leu Val Arg Gln Gly Leu Ala Lys Val Ala Tyr Val Tyr Lys Pro

115 120 125

Glu Phe Met

130

<210> 19

<211> 369

<212> DNA

<213> Artificial Sequence

<220>

<223> bacterial expression vector pMF that encodes a fusion protein, containing a sequence for a gene encoding the N-terminal nucleotide binding domain of *staph. nuclease*, designated SFC120

<220>

<221> CDS

<222> (1)...(369)

<400> 19

aca gcc atg gca act tca act aaa aaa tta cat aaa gaa cct gcg act 48

Thr Ala Met Ala Thr Ser Thr Lys Lys Leu His Lys Glu Pro Ala Thr

1 5 10 15

tta att aaa gcg att gat ggt gat acg gtt aaa tta ttg tac aaa ggt 96

Leu Ile Lys Ala Ile Asp Gly Asp Thr Val Lys Leu Leu Tyr Lys Gly

20 25 30

caa cca ttg aca ttc aga cta tta ttg gtt gat aca cct gaa aca aag 144

Gln Pro Leu Thr Phe Arg Leu Leu Leu Val Asp Thr Pro Glu Thr Lys

35 40 45

cat cct aaa aaa ggt gta gag aaa tat ggt cct gaa gca agt gca ttt 192

His Pro Lys Lys Gly Val Glu Lys Tyr Gly Pro Glu Ala Ser Ala Phe

50 55 60

acg aaa aaa ttg gta gaa aat gca aag aaa att gaa gtc gag ttt gac 240

Thr Lys Lys Leu Val Glu Asn Ala Lys Lys Ile Glu Val Glu Phe Asp

65 70 75 80

aaa ggt caa aga act gat aaa tat gga cgt ggc tta gcg tat att tat 288

Lys Gly Gln Arg Thr Asp Lys Tyr Gly Arg Gly Leu Ala Tyr Ile Tyr

85

90

95

gct gat gga aaa ttg gta aac gaa gct tta gtt cgt caa ggc ttg gct 336

Ala Asp Gly Lys Leu Val Asn Glu Ala Leu Val Arg Gln Gly Leu Ala

100

105

110

aaa gtt gct tat gtt tac aaa cct gaa ttc atg

369

Lys Val Ala Tyr Val Tyr Lys Pro Glu Phe Met

115

120

<210> 20

<211> 123

<212> PRT

<213> Artificial Sequence

<220>

<223> fusion protein MF encoded by bacterial expression vector pMF, containing a sequence for a gene encoding the N-terminal nucleotide binding domain of *staph.* nuclease, designated SFC120

<400> 20

Thr Ala Met Ala Thr Ser Thr Lys Lys Leu His Lys Glu Pro Ala Thr

1

5

10

15

Leu Ile Lys Ala Ile Asp Gly Asp Thr Val Lys Leu Leu Tyr Lys Gly

20 25 30
Gln Pro Leu Thr Phe Arg Leu Leu Leu Val Asp Thr Pro Glu Thr Lys
35 40 45
His Pro Lys Lys Gly Val Glu Lys Tyr Gly Pro Glu Ala Ser Ala Phe
50 55 60
Thr Lys Lys Leu Val Glu Asn Ala Lys Lys Ile Glu Val Glu Phe Asp
65 70 75 80
Lys Gly Gln Arg Thr Asp Lys Tyr Gly Arg Gly Leu Ala Tyr Ile Tyr
85 90 95
Ala Asp Gly Lys Leu Val Asn Glu Ala Leu Val Arg Gln Gly Leu Ala
100 105 110
Lys Val Ala Tyr Val Tyr Lys Pro Glu Phe Met
115 120

<210> 21

<211> 390

<212> DNA

<213> Artificial Sequence

<220>

<223> bacterial expression vector pMFH that encodes a fusion protein, containing a sequence for a gene encoding the N-terminal nucleotide binding domain of staph. nuclease, designated SFC120

<220>

<221> CDS

<222> (1) ... (390)

<400> 21

aca gcc atg gca act tca act aaa aaa tta cat aaa gaa cct gcg act 48

Thr Ala Met Ala Thr Ser Thr Lys Lys Leu His Lys Glu Pro Ala Thr

1

5

10

15

tta att aaa gcg att gat ggt gat acg gtt aaa tta ttg tac aaa ggt 96

Leu Ile Lys Ala Ile Asp Gly Asp Thr Val Lys Leu Leu Tyr Lys Gly

20

25

30

caa cca ttg aca ttc aga cta tta ttg gtt gat aca cct gaa aca aag 144

Gln Pro Leu Thr Phe Arg Leu Leu Leu Val Asp Thr Pro Glu Thr Lys

35

40

45

cat cct aaa aaa ggt gta gag aaa tat ggt cct gaa gca agt gca ttt 192

His Pro Lys Lys Gly Val Glu Lys Tyr Gly Pro Glu Ala Ser Ala Phe

50

55

60

acg aaa aaa ttg gta gaa aat gca aag aaa att gaa gtc gag ttt gac 240

Thr Lys Lys Leu Val Glu Asn Ala Lys Lys Ile Glu Val Glu Phe Asp

65

70

75

80

aaa ggt caa aga act gat aaa tat gga cgt ggc tta gcg tat att tat 288

Lys Gly Gln Arg Thr Asp Lys Tyr Gly Arg Gly Leu Ala Tyr Ile Tyr

85

90

95

gct gat gga aaa ttg gta aac gaa gct tta gtt cgt caa ggc ttg gct 336

Ala Asp Gly Lys Leu Val Asn Glu Ala Leu Val Arg Gln Gly Leu Ala

100

105

110

aaa gtt gct tat gtt tac aaa cct cac cat cac cat cac cat atg gaa 384

Lys Val Ala Tyr Val Tyr Lys Pro His His His His His His Met Glu

115

120

125

ttc atg

390

Phe Met

130

<210> 22

<211> 130

<212> PRT

<213> Artificial Sequence

<220>

<223> fusion protein MFH encoded by bacterial expression vector pMFH, containing a sequence for a gene encoding the N-terminal nucleotide binding domain of *staph.* nuclease, designated SFC120

<400> 22

Thr Ala Met Ala Thr Ser Thr Lys Lys Leu His Lys Glu Pro Ala Thr

1

5

10

15

Leu Ile Lys Ala Ile Asp Gly Asp Thr Val Lys Leu Leu Tyr Lys Gly

20

25

30

Gln Pro Leu Thr Phe Arg Leu Leu Leu Val Asp Thr Pro Glu Thr Lys

35

40

45

His Pro Lys Lys Gly Val Glu Lys Tyr Gly Pro Glu Ala Ser Ala Phe

50

55

60

Thr Lys Lys Leu Val Glu Asn Ala Lys Lys Ile Glu Val Glu Phe Asp

65

70

75

80

Lys Gly Gln Arg Thr Asp Lys Tyr Gly Arg Gly Leu Ala Tyr Ile Tyr

85

90

95

Ala Asp Gly Lys Leu Val Asn Glu Ala Leu Val Arg Gln Gly Leu Ala

100

105

110

Lys Val Ala Tyr Val Tyr Lys Pro His His His His His His Met Glu

115

120

125

Phe Met

130

<210> 23

<400> 23

000

<210> 24

<400> 24

000

<210> 25

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Forward oligonucleotide primer

<400> 25

catgccatgg gtttcaccca tcaccatcac catgcaactt caactaaa

48

<210> 26

<211> 54

<212> DNA

<213> Artificial Sequence

<220>

<223> Reverse oligonucleotide primer

<400> 26

ggaagatctt tagaattccg cggatccacg cggcttaaatt tgacctgaat cagc

54

<210> 27

<211> 11

<212> PRT

<213> Artificial Sequence

<220>

<223> HCR1 peptide derived from the C-terminal 11 amino acid residues of hirudin and strongly inhibiting thrombin

<400> 27

Asp Phe Glu Glu Ile Pro Glu Glu Tyr Leu Gln

1

5

10

<210> 28

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic complementary oligonucleotide

<400> 28

aattcatggg tgacttcgaa gaaatcccgg aagaatacct gcagtaag

48

<210> 29

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic complementary oligonucleotide

<400> 29

gacccctact gcaggtattc ttccgggatt tcttcgaagt cacccatg

48

<210> 30

<211> 4

<212> PRT

<213> Artificial Sequence

<220>

<223> short peptide substrates for the thrombin active
site

<400> 30

Leu Asp Pro Arg

1

<210> 31

<211> 4

<212> PRT

<213> Artificial Sequence

<220>

<223> short peptide substrates for the thrombin active
site

<400> 31

Val Asp Pro Arg

1

<210> 32

<211> 4

<212> PRT

<213> Artificial Sequence

<220>

<223> short peptide substrates for the thrombin active
site

<400> 32

Phe Asn Pro Arg

1

<210> 33

<211> 4

<212> PRT

<213> Artificial Sequence

<220>

<223> short peptide substrates for the thrombin active
site

<400> 33

Pro Asn Pro Arg

1

<210> 34

<211> 4

<212> PRT

<213> Artificial Sequence

<220>

<223> short peptide substrates for the thrombin active
site

<400> 34

Phe Ser Ala Arg

1

<210> 35

<211> 4

<212> PRT

<213> Artificial Sequence

<220>

<223> short peptide substrates for the thrombin active
site

<400> 35

Val Ser Pro Arg

1

<210> 36

<211> 36

<212> PRT

<213> Artificial Sequence

<220>

<223> Recombinant Thrombin Target Peptide

<400> 36

Met Gly Leu Asp Pro Arg Met Gly Val Asp Pro Arg Met Gly Phe Asn

1 5 10 15

Pro Arg Met Gly Pro Asn Pro Arg Met Gly Phe Ser Pro Arg Met Gly

20 25 30

Val Ser Pro Arg

35

<210> 37

<211> 63

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic oligonucleotides

<400> 37

aattcatggg tctggatcct agaatgggag tagaccacg tatgggcttc aaccgcgca

60

tga

63

<210> 38

<211> 63

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic oligonucleotides

<400> 38

gatctatggg tcctaaccga cgtatgggat ttagcgctcg catgggcgtt tctccacgtt 60

aat 63

<210> 39

<211> 22

<212> PRT

<213> Artificial Sequence

<220>

<223> FD22 peptide refers to an amino acid sequence of 22 residues which contains one thrombin substrate sequence linked to the HRC1 peptide by a native linking sequence of hirudin

<400> 39

Phe Asp Pro Arg Pro Gln Ser His Asn Asp Gly Asp Phe Glu Glu Ile

1

5

10

15

Pro Glu Glu Tyr Leu Gln

20

<210> 40

<211> 78

<212> DNA

<213> Artificial Sequence

<220>

<223> complementary oligonucleotide primer

<400> 40

aattcatgtt tgaccgcgc cctcaaagtc ataacgacgg tgattttgag gaaattcctg 60

aagagtattt acaataag 78

<210> 41

<211> 33

<212> PRT

<213> Artificial Sequence

<220>

<223> EB2 or EB3 peptide refers to cytoplasmic carboxyl-terminal 33 amino acid residue sequence which are conserved among ephrin B1, ephrin B2 and ephrin B3. This particular peptide is responsible for binding to downstream partners such as Grb4 and RGS3 proteins

<220>

<221> VARIANT

<222> 20

<223> Xaa = Glu or Asp

<220>

<221> VARIANT

<222> 21

<223> Xaa = Met or Gly

<220>

<221> VARIANT

<222> 27

<223> Xaa = Ala or Pro

<400> 41

Cys Pro His Tyr Glu Lys Val Ser Gly Asp Tyr Gly His Pro Val Tyr

1

5

10

15

Ile Val Gln Xaa Xaa Pro Pro Gln Ser Pro Xaa Asn Ile Tyr Tyr Lys

20

25

30

Val

<210> 42

<211> 57

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic oligonucleotide primer

<400> 42

ccggaattca tgtgcccgca ctacgaaaaa gtttccggtg actatggtca tccggtt 57

<210> 43

<211> 66

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic oligonucleotide primer

<400> 43

cgcggtatcct taaactttgt agtagatggt tggcgggctc tgtggcggac catcctgaac 60

gatata 66

<210> 44

<211> 48

<212> PRT

<213> Artificial Sequence

<220>

<223> Candida Cla4 protein

<400> 44

Ser Val Leu Thr Gly Gly Asn Ser Gly Val Ser Gly Pro Ile Asn Phe

1

5

10

15

Thr His Lys Val His Val Gly Phe Asp Pro Ala Ser Gly Asn Phe Thr

20

25

30

Gly Leu Pro Asp Thr Trp Lys Ser Leu Leu Gln His Ser Lys Ile Thr

35

40

45

<210> 45

<211> 44

<212> PRT

<213> Artificial Sequence

<220>

<221>

<222>

<223> Candida Ste20 protein

<400> 45

Glu Val Asn Ile Lys Ile Ser Thr Pro Phe Asn Ala Lys His Leu Ala

1

5

10

15

His Val Gly Ile Asp Xaa Asp Asn Gly Ser Tyr Thr Gly Leu Pro Ile

20

25

30

Glu Trp Glu Arg Leu Leu Ser Ala Ser Gly Ile Thr

35

40

<210> 46

<211> 186

<212> DNA

<213> Artificial Sequence

<220>

<223> propeptide of human cathepsin B

<220>

<221> CDS

<222> (1)...(186)

<400> 46

cgg agc agg ccc tct ttc cat ccc gtg tcg gat gag ctg gtc aac tat 48

Arg Ser Arg Pro Ser Phe His Pro Val Ser Asp Glu Leu Val Asn Tyr

1

5

10

15

gtc aac aaa cgg aat acc acg tgg cag gcc ggg cac aac ttc tac aac 96

Val Asn Lys Arg Asn Thr Thr Trp Gln Ala Gly His Asn Phe Tyr Asn

20

25

30

gtg gac atg agc tac ttg aag agg cta tgt ggt acc ttc ctg ggt ggg 144

Val Asp Met Ser Tyr Leu Lys Arg Leu Cys Gly Thr Phe Leu Gly Gly

35

40

45

ccc aag cca ccc cag aga gtt atg ttt acc gag gac ctg aag 186

Pro Lys Pro Pro Gln Arg Val Met Phe Thr Glu Asp Leu Lys

50

55

60

<210> 47

<211> 62

<212> PRT

<213> Artificial Sequence

<220>

<223> propeptide of human cathepsin B

<400> 47

Arg Ser Arg Pro Ser Phe His Pro Val Ser Asp Glu Leu Val Asn Tyr

1 5 10 15

Val Asn Lys Arg Asn Thr Thr Trp Gln Ala Gly His Asn Phe Tyr Asn

20 25 30

Val Asp Met Ser Tyr Leu Lys Arg Leu Cys Gly Thr Phe Leu Gly Gly

35 40 45

Pro Lys Pro Pro Gln Arg Val Met Phe Thr Glu Asp Leu Lys

50 55 60

<210> 48

<211> 14

<212> PRT

<213> Artificial Sequence

<220>

<223> CRIB concensus which comprise the CRIB motif

<220>

<221> VARIANT

<222> (1) ... (14)

<223> Xaa = any amino acid

<400> 48

Ile Ser Xaa Pro Xaa Xaa Phe Xaa His Xaa Xaa His Val Gly

1

5

10